SEQUENCE LISTING

<110> WHITELAW, CHRISTOPHER BRUCE ALEXANDER
CLARK, ANTHONY JOHN
WOLF, CHARLES ROLAND

<120> MULTI-REPORTER GENE MODEL FOR TOXICOLOGICAL SCREENING

<130> 102286.155 US1

<140> 10/522,356
<141> 2005-01-26

<150> PCT/GB03/003192
<151> 2003-07-25

<150> GB 0217402.7
<151> 2002-07-26

<160> 41

<170> PatentIn version 3.3

<210> 1
<211> 10
<212> PRT
<213> Homo sapiens

<400> 1
Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1 5 10

<210> 2
<211> 14
<212> PRT
<213> Simian virus

<400> 2
Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr
1 5 10

<210> 3
<211> 9
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: Haemaglutinin epitope from unknown organism

<400> 3
Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5

<210> 4
<211> 12
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: Clone100
epitope from unknown organism

<400> 4
Asn Val Arg Phe Ser Thr Ile Val Arg Arg Arg Ala
1 5 10

<210> 5
<211> 14
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: Rab11a
epitope from unknown organism

<400> 5
Lys Gln Met Ser Asp Arg Arg Glu Asn Asp Met Ser Pro Ser
1 5 10

<210> 6
<211> 15
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: DOB
epitope from unknown organism

<400> 6
Ser Gly Asn Glu Val Ser Arg Ala Val Leu Leu Pro Gln Ser Cys
1 5 10 15

<210> 7
<211> 17
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: SG11
epitope from unknown organism

<400> 7
Ser Ser Leu Ser Tyr Thr Asn Pro Ala Val Ala Ala Thr Ser Ala Asn
1 5 10 15

Leu

<210> 8
<211> 15
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: ErbB4
epitope from unknown organism

<400> 8
Arg Ser Thr Leu Gln His Pro Asp Tyr Leu Gln Glu Tyr Ser Thr
1 5 10 15

<210> 9
<211> 16
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: ARF
epitope from unknown organism

<400> 9
Val Ser Thr Leu Leu Arg Trp Glu Arg Phe Pro Gly His Arg Gln Ala
1 5 10 15

<210> 10
<211> 20
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: RYK
epitope from unknown organism

<400> 10
Lys Phe Gln Gln Leu Val Gln Cys Leu Thr Glu Phe His Ala Ala Leu
1 5 10 15

Gly Ala Tyr Val
20

<210> 11
<211> 20
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: WILPEP1
epitope from unknown organism

<400> 11
Gln Glu Gln Cys Gln Glu Val Trp Arg Lys Arg Val Ile Ser Ala Phe
1 5 10 15

Leu Lys Ser Pro
20

<210> 12
<211> 14
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown Sequence: HAF10
epitope from unknown organism

<400> 12
Arg Leu Ser Asp Lys Thr Gly Pro Val Ala Gln Glu Lys Ser
1 5 10

<210> 13
<211> 543
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (1)..(540)

<400> 13
atg aag atg ctg ctg ctg ctg tgt ttg gga ctg acc cta gtc tgt gtc 48
Met Lys Met Leu Leu Leu Leu Cys Leu Gly Leu Thr Leu Val Cys Val
1 5 10 15

cat gca gaa gaa gct agt tct acg gga agg aac ttt aat gta gaa aag 96
His Ala Glu Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val Glu Lys
20 25 30

att aat ggg gaa tgg cat act att atc ctg gcc tct gac aaa aga gaa 144
Ile Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys Arg Glu
35 40 45

aag ata gaa gat aat ggc aac ttt aga ctt ttt ctg gag caa atc cat 192
Lys Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln Ile His
50 55 60

gtc ttg gag aaa tcc tta gtt ctt aaa ttc cat act gta aga gat gaa 240
Val Leu Glu Lys Ser Leu Val Leu Lys Phe His Thr Val Arg Asp Glu
65 70 75 80

gag tgc tcg gaa tta tct atg gtt gct gac aaa aca gaa aag gct ggt 288
Glu Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys Ala Gly
85 90 95

```
gaa tat tct gtg acg tat gat gga ttc aat aca ttt act ata cct aag      336
Glu Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile Pro Lys
            100                 105                 110

aca gac tat gat aac ttt ctt atg gct cat ctc att aac gaa aag gat      384
Thr Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu Lys Asp
        115                 120                 125

ggg gaa acc ttc cag ctg atg ggg ctc tat ggc cga gaa cca gat ttg      432
Gly Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu Pro Asp Leu
    130                 135                 140

agt tca gac atc aag gaa agg ttt gca caa cta tgt gag aag cat gga      480
Ser Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu Lys His Gly
145                 150                 155                 160

atc ctt aga gaa aat atc att gac cta tcc aat gcc aat cgc tgc ctc      528
Ile Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn Arg Cys Leu
                165                 170                 175

cag gcc cga gaa tga                                                 543
Gln Ala Arg Glu
            180
```

<210> 14
<211> 180
<212> PRT
<213> Mus musculus

<400> 14

```
Met Lys Met Leu Leu Leu Leu Cys Leu Gly Leu Thr Leu Val Cys Val
1               5                   10                  15

His Ala Glu Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val Glu Lys
            20                  25                  30

Ile Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys Arg Glu
        35                  40                  45

Lys Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln Ile His
    50                  55                  60

Val Leu Glu Lys Ser Leu Val Leu Lys Phe His Thr Val Arg Asp Glu
65                  70                  75                  80

Glu Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys Ala Gly
                85                  90                  95

Glu Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile Pro Lys
            100                 105                 110
```

```
Thr Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu Lys Asp
        115                 120                 125

Gly Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu Pro Asp Leu
    130                 135                 140

Ser Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu Lys His Gly
145                 150                 155                 160

Ile Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn Arg Cys Leu
                165                 170                 175

Gln Ala Arg Glu
            180
```

<210> 15
<211> 540
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant mMUP reporter molecule

<220>
<221> CDS
<222> (1)..(537)

<400> 15

```
ggg ccc ctg gga tcc atg gag cag aaa ctc atc tct gaa gag gat ctg       48
Gly Pro Leu Gly Ser Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1               5                   10                  15

acc atg gaa gct agt tct acg gga agg aac ttt aat gta gaa aag att       96
Thr Met Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val Glu Lys Ile
            20                  25                  30

aat ggg gaa tgg cat act att atc ctg gcc tct gac aaa aga gaa aag      144
Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys Arg Glu Lys
        35                  40                  45

ata gaa gat aat ggc aac ttt aga ctt ttt ctg gag caa atc cat gtc      192
Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln Ile His Val
    50                  55                  60

ttg gag aaa tcc tta gtt ctt aaa ttc cat act gta aga gat gaa gag      240
Leu Glu Lys Ser Leu Val Leu Lys Phe His Thr Val Arg Asp Glu Glu
65                  70                  75                  80

tgc tcg gaa tta tct atg gtt gct gac aaa aca gaa aag gct ggt gaa      288
Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys Ala Gly Glu
                85                  90                  95
```

```
tat tct gtg acg tat gat gga ttc aat aca ttt act ata cct aag aca      336
Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile Pro Lys Thr
            100                 105                 110

gac tat gat aac ttt ctt atg gct cat ctc att aac gaa aag gat ggg      384
Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu Lys Asp Gly
        115                 120                 125

gaa acc ttc cag ctg atg ggg ctc tat ggc cga gaa cca gat ttg agt      432
Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu Pro Asp Leu Ser
    130                 135                 140

tca gac atc aag gaa agg ttt gca caa cta tgt gag aag cat gga atc      480
Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu Lys His Gly Ile
145                 150                 155                 160

ctt aga gaa aat atc att gac cta tcc aat gcc aat cgc tgc ctc cag      528
Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn Arg Cys Leu Gln
                165                 170                 175

gcc cga gaa tga                                                     540
Ala Arg Glu
```

<210> 16
<211> 179
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant mMUP reporter molecule

<400> 16

```
Gly Pro Leu Gly Ser Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1               5                   10                  15

Thr Met Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val Glu Lys Ile
            20                  25                  30

Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys Arg Glu Lys
        35                  40                  45

Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln Ile His Val
    50                  55                  60

Leu Glu Lys Ser Leu Val Leu Lys Phe His Thr Val Arg Asp Glu Glu
65                  70                  75                  80

Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys Ala Gly Glu
                85                  90                  95
```

```
Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile Pro Lys Thr
            100                 105                 110

Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu Lys Asp Gly
        115                 120                 125

Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu Pro Asp Leu Ser
    130                 135                 140

Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu Lys His Gly Ile
145                 150                 155                 160

Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn Arg Cys Leu Gln
                165                 170                 175

Ala Arg Glu
```

<210> 17
<211> 540
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant BLGm reporter molecule

<220>
<221> CDS
<222> (1)..(537)

<400> 17

```
ggg ccc ctg gga tcc atg gcc atc atc gtc acc cag acc atg aaa ggc       48
Gly Pro Leu Gly Ser Met Ala Ile Ile Val Thr Gln Thr Met Lys Gly
1               5                   10                  15

ctg gac atc cag aag gtg gcg ggg act tgg cac tcc ttg gct atg gcg       96
Leu Asp Ile Gln Lys Val Ala Gly Thr Trp His Ser Leu Ala Met Ala
            20                  25                  30

gcc agc gac atc tcc ctg ctg gat gcc cag agt gcc ccc ctg aga gtg      144
Ala Ser Asp Ile Ser Leu Leu Asp Ala Gln Ser Ala Pro Leu Arg Val
        35                  40                  45

tac gtg gag gag ctg aag ccc acc ccc gag ggc aac ctg gag atc ctg      192
Tyr Val Glu Glu Leu Lys Pro Thr Pro Glu Gly Asn Leu Glu Ile Leu
    50                  55                  60
```

```
ctg cag aaa tgg gag aac ggc gag tgt gct cag aag aag att att gca      240
Leu Gln Lys Trp Glu Asn Gly Glu Cys Ala Gln Lys Lys Ile Ile Ala
65                  70                  75                  80

gaa aaa acc aag atc cct gcg gtg ttc aag atc gat gcc ttg aat gag      288
Glu Lys Thr Lys Ile Pro Ala Val Phe Lys Ile Asp Ala Leu Asn Glu
                85                  90                  95

aac aaa gtc ctt gtg ctg gac acc gac tac aaa aag tac ctg ctc ttc      336
Asn Lys Val Leu Val Leu Asp Thr Asp Tyr Lys Lys Tyr Leu Leu Phe
            100                 105                 110

tgc atg gaa aac agt gct gag ccc gag caa agc ctg gcc tgc cag tgc      384
Cys Met Glu Asn Ser Ala Glu Pro Glu Gln Ser Leu Ala Cys Gln Cys
        115                 120                 125

ctg gtc agg acc ccg gag gtg gac aac gag gcc ctg gag aaa ttc gac      432
Leu Val Arg Thr Pro Glu Val Asp Asn Glu Ala Leu Glu Lys Phe Asp
    130                 135                 140

aaa gcc ctc aag gcc ctg ccc atg cac atc cgg ctt gcc ttc aac ccg      480
Lys Ala Leu Lys Ala Leu Pro Met His Ile Arg Leu Ala Phe Asn Pro
145                 150                 155                 160

acc cag ctg gag ggg cag tgc cac gtc gag cag aaa ctc atc tct gaa      528
Thr Gln Leu Glu Gly Gln Cys His Val Glu Gln Lys Leu Ile Ser Glu
                165                 170                 175

gag gat ctg tag                                                     540
Glu Asp Leu
```

<210> 18
<211> 179
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant BLGm reporter molecule

<400> 18

```
Gly Pro Leu Gly Ser Met Ala Ile Ile Val Thr Gln Thr Met Lys Gly
1               5                   10                  15

Leu Asp Ile Gln Lys Val Ala Gly Thr Trp His Ser Leu Ala Met Ala
            20                  25                  30

Ala Ser Asp Ile Ser Leu Leu Asp Ala Gln Ser Ala Pro Leu Arg Val
        35                  40                  45

Tyr Val Glu Glu Leu Lys Pro Thr Pro Glu Gly Asn Leu Glu Ile Leu
    50                  55                  60
```

Leu Gln Lys Trp Glu Asn Gly Glu Cys Ala Gln Lys Lys Ile Ile Ala
65 70 75 80

Glu Lys Thr Lys Ile Pro Ala Val Phe Lys Ile Asp Ala Leu Asn Glu
85 90 95

Asn Lys Val Leu Val Leu Asp Thr Asp Tyr Lys Lys Tyr Leu Leu Phe
100 105 110

Cys Met Glu Asn Ser Ala Glu Pro Glu Gln Ser Leu Ala Cys Gln Cys
115 120 125

Leu Val Arg Thr Pro Glu Val Asp Asn Glu Ala Leu Glu Lys Phe Asp
130 135 140

Lys Ala Leu Lys Ala Leu Pro Met His Ile Arg Leu Ala Phe Asn Pro
145 150 155 160

Thr Gln Leu Glu Gly Gln Cys His Val Glu Gln Lys Leu Ile Ser Glu
165 170 175

Glu Asp Leu

<210> 19
<211> 214
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic modified MUP protein produced from the pSecTag vector

<400> 19
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15

Gly Ser Thr Gly Asp Ala Ala Gln Pro Ala Arg Arg Ala Arg Arg Thr
20 25 30

Lys Leu Gly Thr Glu Leu Gly Ser Met Glu Gln Lys Leu Ile Ser Glu
35 40 45

Glu Asp Leu Thr Met Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val
50 55 60

```
Glu Lys Ile Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys
65                  70                  75                  80

Arg Glu Lys Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln
                85                  90                  95

Ile His Val Leu Glu Lys Ser Leu Val Leu Lys Phe His Thr Val Arg
            100                 105                 110

Asp Glu Glu Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys
        115                 120                 125

Ala Gly Glu Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile
    130                 135                 140

Pro Lys Thr Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu
145                 150                 155                 160

Lys Asp Gly Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu Pro
                165                 170                 175

Asp Leu Ser Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu Lys
            180                 185                 190

His Gly Ile Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn Arg
        195                 200                 205

Cys Leu Gln Ala Arg Glu
    210
```

<210> 20
<211> 243
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic modified MUP protein produced from the pSecTag vector

<400> 20

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                   10                  15
```

```
Gly Ser Thr Gly Asp Ala Ala Gln Pro Ala Lys Met Leu Leu Leu Leu
            20                  25                  30

Cys Leu Gly Leu Thr Leu Val Cys Val His Ala Glu Glu Ala Ser Ser
        35                  40                  45

Thr Gly Arg Asn Phe Asn Val Glu Lys Ile Asn Gly Glu Trp His Thr
    50                  55                  60

Ile Ile Leu Ala Ser Asp Lys Arg Glu Lys Ile Glu Asp Asn Gly Asn
65                  70                  75                  80

Phe Arg Leu Phe Leu Glu Gln Ile His Val Leu Glu Lys Ser Leu Val
                85                  90                  95

Leu Lys Phe His Thr Val Arg Asp Glu Glu Cys Ser Glu Leu Ser Met
            100                 105                 110

Val Ala Asp Lys Thr Glu Lys Ala Gly Glu Tyr Ser Val Thr Tyr Asp
        115                 120                 125

Gly Phe Asn Thr Phe Thr Ile Pro Lys Thr Asp Tyr Asp Asn Phe Leu
    130                 135                 140

Met Ala His Leu Ile Asn Glu Lys Asp Gly Glu Thr Phe Gln Leu Met
145                 150                 155                 160

Gly Leu Tyr Gly Arg Glu Pro Asp Leu Ser Ser Asp Ile Lys Glu Arg
                165                 170                 175

Phe Ala Gln Leu Cys Glu Lys His Gly Ile Leu Arg Glu Asn Ile Ile
            180                 185                 190

Asp Leu Ser Asn Ala Asn Arg Cys Leu Gln Ala Arg Glu Glu Gln Lys
        195                 200                 205

Leu Ile Ser Glu Glu Asp Leu Ala Ala Ala Arg Gly Gly Pro Glu Gln
    210                 215                 220

Lys Leu Ile Ser Glu Glu Asp Leu Asn Ser Ala Val Asp His His His
225                 230                 235                 240

His His His
```

<210> 21
<211> 253
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic modified MUP protein produced from the pSecTag vector

<400> 21

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                  10                  15

Gly Ser Thr Gly Asp Ala Ala Gln Pro Ala Lys Met Leu Leu Leu Leu
            20                  25                  30

Cys Leu Gly Leu Thr Leu Val Cys Val His Ala Glu Glu Ala Ser Ser
        35                  40                  45

Thr Gly Arg Asn Phe Asn Val Glu Lys Ile Asn Gly Glu Trp His Thr
    50                  55                  60

Ile Ile Leu Ala Ser Asp Lys Arg Glu Lys Ile Glu Asp Asn Gly Asn
65                  70                  75                  80

Phe Arg Leu Phe Leu Glu Gln Ile His Val Leu Glu Lys Ser Leu Val
                85                  90                  95

Leu Lys Phe His Thr Val Arg Asp Glu Glu Cys Ser Glu Leu Ser Met
            100                 105                 110

Val Ala Asp Lys Thr Glu Lys Ala Gly Glu Tyr Ser Val Thr Tyr Asp
        115                 120                 125

Gly Phe Asn Thr Phe Thr Ile Pro Lys Thr Asp Tyr Asp Lys Leu Gly
    130                 135                 140

Thr Gly Ser Ser Ser Glu Phe Asn Phe Leu Met Ala His Leu Ile Asn
145                 150                 155                 160

Glu Lys Asp Gly Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu
                165                 170                 175
```

```
Pro Asp Leu Ser Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu
            180                 185                 190

Lys His Gly Ile Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn
        195                 200                 205

Arg Cys Leu Gln Ala Arg Glu Glu Gln Lys Leu Ile Ser Glu Glu Asp
    210                 215                 220

Leu Ala Ala Ala Arg Gly Gly Pro Glu Gln Lys Leu Ile Ser Glu Glu
225                 230                 235                 240

Asp Leu Asn Ser Ala Val Asp His His His His His His
                245                 250
```

<210> 22
<211> 259
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic modified MUP protein produced from the pSecTag vector

<400> 22

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                   10                  15

Gly Ser Thr Gly Asp Ala Ala Gln Pro Ala Lys Met Leu Leu Leu Leu
            20                  25                  30

Cys Leu Gly Leu Thr Leu Val Cys Val His Ala Glu Glu Ala Ser Ser
        35                  40                  45

Thr Gly Arg Asn Phe Asn Val Glu Lys Ile Asn Gly Glu Trp His Thr
    50                  55                  60

Ile Ile Leu Ala Ser Asp Lys Arg Glu Lys Ile Glu Asp Asn Gly Asn
65                  70                  75                  80

Phe Arg Leu Phe Leu Glu Gln Ile His Val Leu Glu Lys Ser Leu Val
                85                  90                  95

Leu Lys Phe His Thr Val Arg Asp Glu Glu Cys Ser Glu Leu Ser Met
            100                 105                 110
```

Val Ala Asp Lys Thr Glu Lys Ala Gly Glu Tyr Ser Val Thr Tyr Asp
115 120 125

Gly Phe Asn Thr Phe Thr Ile Pro Lys Thr Asp Tyr Asp Lys Leu Asn
130 135 140

Val Arg Phe Ser Thr Ile Val Arg Arg Arg Ala Glu Phe Asn Phe Leu
145 150 155 160

Met Ala His Leu Ile Asn Glu Lys Asp Gly Glu Thr Phe Gln Leu Met
165 170 175

Gly Leu Tyr Gly Arg Glu Pro Asp Leu Ser Ser Asp Ile Lys Glu Arg
180 185 190

Phe Ala Gln Leu Cys Glu Lys His Gly Ile Leu Arg Glu Asn Ile Ile
195 200 205

Asp Leu Ser Asn Ala Asn Arg Cys Leu Gln Ala Arg Glu Glu Gln Lys
210 215 220

Leu Ile Ser Glu Glu Asp Leu Ala Ala Ala Arg Gly Gly Pro Glu Gln
225 230 235 240

Lys Leu Ile Ser Glu Glu Asp Leu Asn Ser Ala Val Asp His His His
245 250 255

His His His

<210> 23
<211> 7379
<212> DNA
<213> Ovis aries

<400> 23
gtgctcagca acacacccag caccagcatt cccgctgctc ctgaggtctg caggcagctc 60

gctgtagcct gagcggtgtg gagggaagtg tcctgggaga tttaaaatgt gagaggcggg 120

aggtgggagg ttgggccctg tgggcctgcc catcccacgt gcctgcatta gccccagtgc 180

tgctcagccg tgcccccgcc gcaggggtca ggtcactttc ccgtcctggg gttattatga 240

ctcttgtcat tgccattgcc atttttgcta ccctaactgg gcagcaggtg cttgcagagc 300

cctcgatacc gaccaggtcc tccctcggag ctcgacctga accccatgtc acccttgccc 360

cagcctgcag agggtgggtg actgcagaga tcccttcacc caaggccacg gtcacatggt 420

ttggaggagc tggtgcccaa ggcagaggcc accctccagg acacacctgt ccccagtgct 480

ggctctgacc tgtccttgtc taagaggctg accccggaag tgttcctggc actggcagcc 540

agcctggacc cagagtccag acacccacct gtgcccccgc ttctggggtc taccaggaac 600

cgtctaggcc cagaggggga cttcctgctt ggccttggat ggaagaaggc ctcctattgt 660

cctcgtagag gaagccaccc cggggcctga ggatgagcca agtgggattc cgggaaccgc 720

gtggctgggg gcccagcccg ggctggctgg cctgcatgcg cctcctgtat aaggccccaa 780

gcctgcctgt ctcagccctc cactccctgc agagctcaga agcacgaccc cagctgcagc 840

catgaagtgc ctcctgcttg ccctgggcct ggccctcgcc tgtggcgtcc aggccatcat 900

cgtcacccag accatgaaag gcctggacat ccagaaggtt cgagggttgg ccgggtgggt 960

gagttgcagg gcgggcaggg gagctgggcc tcagagagcc aagagaggct gtgacgttgg 1020

gttcccatca gtcagctagg gccacctgac aaatccccgc tggggcagct tcaaccaggc 1080

gttcactgtc ttgcattctg gaggctggaa gcccaagatc caggtgttgg cagggctggc 1140

ttctcctgcg gccgctctct ggggagcaga cggccgtctt ctccagtcct ctgcgcgccc 1200

tgatttcctc ttcctgtgag gccaccaggc ctgctggaaa cacgcctgcc tgcgcagctt 1260

cacacgacct ttgtcatctc tttaaaggcc atgtctccag agtcatgtgt tgaagttctg 1320

ggggttagtg ggacacagtt cagcccctaa aagagtctct ctgcccctca aattttcccc 1380

acctccagcc atgtctcccc aagatccaaa tgttgctaca tgtgggggg ctcatctggg 1440

tccctctttg ggttcagtgt gagtctgggg agagcattcc ccagggtgca gagttggggg 1500

gagtatctca gggctgccca ggccggggtg ggacagagag cccactgtgg ggctgggggc 1560

cccttcccac ccccagagtg caactcaagg tccctctcca ggtggcgggg acttggcact 1620

ccttggctat ggcggccagc gacatctccc tgctggatgc ccagagtgcc cccctgagag 1680

tgtacgtgga ggagctgaag cccacccccg agggcaacct ggagatcctg ctgcagaaat 1740

ggtgggcgtc tctccccaac atggaacccc cactccccag ggctgtggac cccccggggg 1800

gtggggtgca ggagggacca gggccccagg gctggggaag agggctcaga gtttactggt 1860

acccggcgct ccacccaagg ctgcccaccc agggcttttt ttttttttaa acttttatta 1920

atttgatgct tcagaacatc atcaaacaaa tgaacataaa acattcattt ttgtttactt 1980

ggaaggggag ataaaatcct ctgaagtgga aatgcatagc aaagatacat acaatgaggc 2040

aggtattctg aattccctgt tagtctgagg attacaagtg tatttgagca acagagagac 2100

attttcatca tttctagtct gaacacctca gtatctaaaa tgaacaagaa gtcctggaaa 2160

cgaagcagtg tggggatagg cccgtgtgaa ggctgctggg aggcagcaga cctgggtctt 2220

cgggctcaag cagttcccgc taccagccct gtccacctca gacgggggtc agggtgcagg 2280

agagagctgg atgggtgtgg gggcagagat ggggacctga accccagggc tgccttttgg 2340

gggtgcctgt ggtcaaggct ctccctgacc ttttctctct ggcttcatct gacttctcct 2400

ggcccatcca cccggtcccc tgtggcctga ggtgacagtg agtgcgccga ggctagttgg 2460

ccagctggct cctatgccca tgccaccccc ctccagccct cctgggccag cttctgcccc 2520

tggccctcag ttcatcctga tgaaaatggt ccatgccaat ggctcagaaa gcagctgtct 2580

ttcagggaga acggcgagtg tgctcagaag aagattattg cagaaaaaac caagatccct 2640

gcggtgttca agatcgatgg tgagtccggg tccctggggg acacccacca cccccgcccc 2700

cggggactgt ggacaggttc agggggctgg cgtcgggccc tgggatgcta agggactggt 2760

ggtgatgaag acactgcctt gacacctgct tcacttgcct cccctgccac ctgcccgggg 2820

ccttggggcg gtggccatgg gcaggtcccg gctggcgggc taacccacca gggtgacacc 2880

cgagctctct ttgctggggg gcgggcggtg ctctgggccc tcaggctgag ctcaggaggt 2940

acctgtgccc tcccaggggt aaccgagagc cgttgcccac tccaggggcc caggtgcccc 3000

acgaccccag cccgctccac agctccttca tctcctggag acaaactctg tccgccctcg 3060

ctcattcact tgttcgtcct aaatccgaga tgataaagct tcgagggggg gttggggttc 3120

catcagggct gcccttccgc cgggcagcct gggccacatc tgcccttggc cccctcagga 3180

ctcactctga ctggaggccc tgcactgact gacgccaggg tgcccagccc agggtctctg 3240

gcgccatcca gctgcactgg gtttgggtgc tggtcctgcc cccaagctgc ccggacacca 3300

caggcagccg gggctgccca ctggcctcgg tcagggtgag ccccagctgc ccccgctcag 3360

ggcttgcccc gacaatgacc ccatcctcag gacgcacccc ccttcccttg ctgggcagtg 3420

tccagcccca cccgagatcg ggggaagccc tatttcttga caactccagt ccctggggga 3480

gggggcctca gactgagtgg tgagtgttcc caagtccagg aggtggtgga gggtcctggc 3540

ggatccagag ttgacagtga gggcttcctg ggccccatgc gcctggcagt ggcagcaggg 3600

aagaggaagc accatttcag gggtggggga tgccagaggc gctccccacc ccgtcttcgc 3660

cgggtggtga ccccgggggа gccccgctgg tcgtggaggg tgctgggggc tgactagcaa 3720

cccctccccc cccgttggaa ctcacttttc tcccgtcttg accgcgtcca gccttgaatg 3780

agaacaaagt ccttgtgctg gacaccgact acaaaaagta cctgctcttc tgcatggaaa 3840

acagtgctga gcccgagcaa agcctggcct gccagtgcct gggtgggtgc caaccctggc 3900

tgcccaggga gaccagctgc gtggtccttg ctgcaacagg gggtgggggg tgggagcttg 3960

atccccagga ggaggagggg tggggggtcc ctgagtcccg ccaggagaga gtggtcgcat 4020

accgggagcc agtctgctgt gggcctgtgg gtggctgggg acgggggcca gacacacagg 4080

ccgggagacg ggtgggctgc agaactgtga ctggtgtgac cgtcgcgatg gggccggtgg 4140

tcactgaatc taacagcctt tgttaccggg gagtttcaat tatttcccaa aataagaact 4200

caggtacaaa gccatctttc aactatcaca tcctgaaaac aaatggcagg tgacattttc 4260

tgtgccgtag cagtcccact gggcattttc agggcccctg tgccaggggg gcgcgggcat 4320

cggcgagtgg aggctcctgg ctgtgtcagc cggcccaggg ggaggaaggg acccggacag 4380

ccagaggtgg ggggcaggct ttccccctgt gacctgcaga cccactgcac tgccctggga 4440

ggaagggagg ggaactaggc caaggggggaa gggcaggtgc tctggagggc aagggcagac 4500

ctgcagacca ccctggggag cagggactga cccccgtccc tgccccatag tcaggacccc 4560

ggaggtggac aacgaggccc tggagaaatt cgacaaagcc ctcaaggccc tgcccatgca 4620

catccggctt gccttcaacc cgacccagct ggagggtgag cacccaggcc ccgcccttcc 4680

ccagggcagg agccacccgg ccccgggacg acctcctccc atggtgaccc ccagctcccc 4740

aggcctccca ggaggaaggg gtggggtgca gcaccccgtg ggggcccccc ccccaccccc 4800

tgccaggcct ctcttcccga ggtgtccagt cccatcctga ccccccccatg actctccctc 4860

ccccacaggg cagtgccacg tctaggtgag cccctgccgg tgcctctggg gtaagctgcc 4920

tgccctgccc cacgtcctgg gcacacacat ggggtagggg gtcttggtgg ggcctgggac 4980

cccacatcag gccctggggt ccccccctgtg agaatggctg gaagctgggg tccctcctgg 5040

cgactgcaga gctggctggc cgcgtgccac tcttgtgggt gacctgtgtc ctggcctcac 5100

acactgacct cctccagctc cttccagcag agctaaggct aagtgagcca gaatggtacc 5160

taaggggagg ctagcggtcc ttctcccgag gaggggctgt cctggaacca ccagccatgg 5220

agaggctggc aagggtctgg caggtgcccc aggaatcaca ggggggcccc atgtccattt 5280

cagggcccgg gagccttgga ctcctctggg gacagacgac gtcaccaccg ccccccccc 5340

atcaggggga ctagaaggga ccaggactgc agtcaccctt cctgggaccc aggcccctcc 5400

aggcccctcc tggggctcct gctctgggca gcttctcctt caccaataaa ggcataaacc 5460

tgtgctctcc cttctgagtc tttgctggac gacgggcagg gggtggagaa gtggtgggga 5520

gggagtctgg ctcagaggat gacagcgggg ctgggatcca gggcgtctgc atcacagtct 5580

tgtgacaact gggggcccac acacatcact gcggctcttt gaaactttca ggaaccaggg 5640

agggactcgg cagagacatc tgccagttca cttggagtgt tcagtcaaca cccaaactcg 5700

acaaaggaca gaaagtggaa aatggctgtc tcttagtcta ataaatattg atatgaaact 5760

caagttgctc atggatcaat atgcctttat gatccagcca gccactactg tcgtatcaac 5820

tcatgtaccc aaacgcactg atctgtctgg ctaatgatga gagattccca gtagagagct 5880

ggcaagaggt cacagtgaga actgtctgca cacacagcag agtccaccag tcatcctaag 5940

gagatcagtc ctggtgttca ttggaggact gatgttgaag ctgaaactcc aatgctttgg 6000

ccacctgatg tgaagagctg actcatttga aaagaccctg atgctgggaa agattgaggg 6060

caggaggaga aggggacgac agaggatgag atggttggat ggcatcacca acacaatgga 6120

catgggtttg ggtggactcc aggagttggt gatggacagg gaggcctggc gtgctacgga 6180

agcggtttat ggggtcacaa agactgagtg actgaactga gctgaactga atggaaatga 6240

ggtatacagc aaagtgggga ttttttagat aataagaata tacacataac atagtgtata 6300

ctcatatttt tatgcatacc tgaatgctca gtcactcagt cgtatctgac tctgtgacct 6360

atggaccgta gccttccagg tttcttctgt ccacagaatt ctccaaggca agaatactgg 6420

agtgggtagc catttcctcc tccaggggat cctcccgacc cagggattga accggcatct 6480

cctgtattgg caggtggatt ctttaccact gtgccaccag ggaagcccgt gttactctct 6540

atgtcccact taattaccaa agctgctcca agaaaaagcc cctgtgccct ctgagcttcc 6600

cggcctgcag agggtggtgg gggtagactg tgacctggga acaccctccc gcttcaggac 6660

tcccgggcca cgtgacccac agtcctgcag acagccgggt agctctgctc ttcaaggctc 6720

attatcttta aaaaaaactg aggtctattt tgtgacttcg ctgccgtaac ttctgaacat 6780

ccagtgcgat ggacaggacc tcctccccag gcctcagggg cttcagggag ccagccttca 6840

cctatgagtc accagacact cgggggtggc cccgccttca gggtgctcac agtcttccca 6900

tcgtcctgat caaagagcaa gaccaatgac ttcttaggag caagcagaca cccacaggac 6960

actgaggttc accagagctg agctgtcctt ttgaacctaa agacacacag ctctcgaagg 7020

ttttctcttt aatctggatt taaggcctac ttgcccctca agagggaaga cagtcctgca 7080

tgtccccagg acagccactc ggtggcatcc gaggccactt agtattatct gaccgcaccc 7140

tggaattaat cggtccaaac tggacaaaaa ccttggtggg aagtttcatc ccagaggcct 7200

caaccatcct gctttgacca ccctgcatct ttttttcttt tatgtgtatg catgtatata 7260

tatatatata ttttttttttt tttcattttt tggctgtgct ggctgttcgt tgcagttcgg 7320

tgcgcaggct tctctctagt ttctctctag tcttctctta tcacagagca gtctctaga 7379

<210> 24
<211> 180
<212> PRT
<213> Ovis aries

<400> 24

```
Met Lys Cys Leu Leu Leu Ala Leu Gly Leu Ala Leu Ala Cys Gly Val
1               5                  10                  15

Gln Ala Ile Ile Val Thr Gln Thr Met Lys Gly Leu Asp Ile Gln Lys
            20                  25                  30

Val Ala Gly Thr Trp His Ser Leu Ala Met Ala Ala Ser Asp Ile Ser
        35                  40                  45

Leu Leu Asp Ala Gln Ser Ala Pro Leu Arg Val Tyr Val Glu Glu Leu
    50                  55                  60

Lys Pro Thr Pro Glu Gly Asn Leu Glu Ile Leu Leu Gln Lys Trp Glu
65                  70                  75                  80

Asn Gly Glu Cys Ala Gln Lys Lys Ile Ile Ala Glu Lys Thr Lys Ile
                85                  90                  95

Pro Ala Val Phe Lys Ile Asp Ala Leu Asn Glu Asn Lys Val Leu Val
            100                 105                 110

Leu Asp Thr Asp Tyr Lys Lys Tyr Leu Leu Phe Cys Met Glu Asn Ser
        115                 120                 125

Ala Glu Pro Glu Gln Ser Leu Ala Cys Gln Cys Leu Val Arg Thr Pro
    130                 135                 140

Glu Val Asp Asn Glu Ala Leu Glu Lys Phe Asp Lys Ala Leu Lys Ala
145                 150                 155                 160

Leu Pro Met His Ile Arg Leu Ala Phe Asn Pro Thr Gln Leu Glu Gly
                165                 170                 175

Gln Cys His Val
            180
```

<210> 25
<211> 925
<212> DNA
<213> Mus musculus

<400> 25
ctgaacccag agagtatata agaacaagca aaggggctgg ggagtggagt gtagccacga 60

tcacaagaaa gacgtggtcc tgacagacag acaatcctat tccctaccaa aatgaagatg 120

ctgctgctgc tgtgtttggg actgacccta gtctgtgtcc atgcagaaga agctagttct 180

acgggaagga actttaatgt agaaaagatt aatggggaat ggcatactat tatcctggcc 240

tctgacaaaa gagaaaagat agaagataat ggcaacttta gacttttct ggagcaaatc 300

catgtcttgg agaattcctt agttcttaaa ttccatactg taagagatga agagtgctcg 360

gaattatcta tggttgctga caaaacagaa aaggctggtg aatattctgt gacgtatgat 420

ggattcaata catttactat acctaagaca gactatgata actttcttat ggctcatctc 480

attaacgaaa aggatgggga aaccttccag ctgatggggc tctatggccg agaaccagat 540

ttgagttcag acatcaagga aaggtttgca caactatgtg agaagcatgg aatccttaga 600

gaaaatatca ttgacctatc caatgccaat cgctgcctcc aggcccgaga atgaagaatg 660

gcctgagcct ccagtgttga gtggagactt ctcaccagga ctccaccatc atcccttcct 720

atccatacag catccccagt ataaattctg tgatctgcat tccatcctgt ctcactgaga 780

agtccaattc cagtctatcc acatgttacc taggatacct catcaagaat caaagacttc 840

tttaaatttt tctttgatat acccatgaca atttttcatg aatttcttcc tcttcctgtt 900

caataaatga ttacccttgc actta 925

<210> 26
<211> 180
<212> PRT
<213> Mus musculus

<400> 26
Met Lys Met Leu Leu Leu Leu Cys Leu Gly Leu Thr Leu Val Cys Val
1 5 10 15

His Ala Glu Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val Glu Lys
20 25 30

Ile Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys Arg Glu
35 40 45

Lys Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln Ile His
50 55 60

Val Leu Glu Asn Ser Leu Val Leu Lys Phe His Thr Val Arg Asp Glu
65 70 75 80

Glu Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys Ala Gly
85 90 95

Glu Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile Pro Lys
100 105 110

Thr Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu Lys Asp
115 120 125

Gly Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu Pro Asp Leu
130 135 140

Ser Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu Lys His Gly
145 150 155 160

Ile Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn Arg Cys Leu
165 170 175

Gln Ala Arg Glu
180

<210> 27
<211> 813
<212> DNA
<213> Rattus norvegicus

<400> 27
ctgctgctgc tgtgtctgcg cctgacactg gtctgtggcc atgcagaaga agctagttcc 60

acaagaggga acctcgatgt ggctaagctc aatggggatt ggttttctat tgtcgtggcc 120

tctaacaaaa gagaaaagat agaagagaat ggcagcatga gagtttttat gcagcacatc 180

gatgtcttgg agaattcctt aggcttcaag ttccgtatta aggaaaatgg agagtgcagg 240

gaactatact tggtttccta caaaacgcca gaggatggtg aatattttgt tgagtatgac 300

ggagggaata catttactat acttaagaca gactactaca tatacgtcat gtttcatctc 360

attaatttca agaacgggga aaccttccag ctgatggtgc tctacggcag aacaaaggat 420

ctgagttcag acatcaagga aaagtttgca aaactatgtg aggcgcatgg aatcactagg 480

gacaatatca ttgatctaac caagactgat cgctgtctcc aggcccgagg atgaagaaag 540

gcctgagcct ccagtgctga gtggagactt ctcaccagga ctctagcatc accatttcct 600

gtccatggag catcctgaga caaattctgc gatctgattt ccatcctctg tcacagaaaa 660

gtgcaatcct ggtctctcca gcatcttccc tagttaccca ggacaacaca tcgagaatta 720

aaagctttct taaatttctc ttggccccac ccatgatcat tccgcacaaa tatcttgctc 780

ttgcagttca ataaatgatt acccttgcac ttt 813

<210> 28
<211> 735
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
GST coding sequence derived from pGEX6p-1

<400> 28
atgtccccta tactaggtta ttggaaaatt aagggccttg tgcaacccac tcgacttctt 60

ttggaatatc ttgaagaaaa atatgaagag catttgtatg agcgcgatga aggtgataaa 120

tggcgaaaca aaaagtttga attgggtttg gagtttccca atcttcctta ttatattgat 180

ggtgatgtta aattaacaca gtctatggcc atcatacgtt atatagctga caagcacaac 240

atgttgggtg gttgtccaaa agagcgtgca gagatttcaa tgcttgaagg agcggttttg 300

gatattagat acggtgtttc gagaattgca tatagtaaag actttgaaac tctcaaagtt 360

gattttctta gcaagctacc tgaaatgctg aaaatgttcg aagatcgttt atgtcataaa 420

acatatttaa atggtgatca tgtaacccat cctgacttca tgttgtatga cgctcttgat 480

gttgttttat acatggaccc aatgtgcctg gatgcgttcc caaaattagt ttgttttaaa 540

aaacgtattg aagctatccc acaaattgat aagtacttga aatccagcaa gtatatagca 600

tggcctttgc agggctggca agccacgttt ggtggtggcg accatcctcc aaaatcggat 660

ctggaagttc tgttccaggg gcccctggga tccccggaat tcccgggtcg actcgagcgg 720

ccgcatcgtg actga 735

<210> 29
<211> 687
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
GST coding sequence derived from pGEX6p-1

<220>
<221> CDS
<222> (1)..(687)

<400> 29

```
atg tcc cct ata cta ggt tat tgg aaa att aag ggc ctt gtg caa ccc       48
Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro
1               5                   10                  15

act cga ctt ctt ttg gaa tat ctt gaa gaa aaa tat gaa gag cat ttg       96
Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu
            20                  25                  30

tat gag cgc gat gaa ggt gat aaa tgg cga aac aaa aag ttt gaa ttg      144
Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu
        35                  40                  45

ggt ttg gag ttt ccc aat ctt cct tat tat att gat ggt gat gtt aaa      192
Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys
    50                  55                  60

tta aca cag tct atg gcc atc ata cgt tat ata gct gac aag cac aac      240
Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn
65                  70                  75                  80

atg ttg ggt ggt tgt cca aaa gag cgt gca gag att tca atg ctt gaa      288
Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu
                85                  90                  95

gga gcg gtt ttg gat att aga tac ggt gtt tcg aga att gca tat agt      336
Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser
            100                 105                 110

aaa gac ttt gaa act ctc aaa gtt gat ttt ctt agc aag cta cct gaa      384
Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu
        115                 120                 125

atg ctg aaa atg ttc gaa gat cgt tta tgt cat aaa aca tat tta aat      432
Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn
    130                 135                 140

ggt gat cat gta acc cat cct gac ttc atg ttg tat gac gct ctt gat      480
Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp
145                 150                 155                 160

gtt gtt tta tac atg gac cca atg tgc ctg gat gcg ttc cca aaa tta      528
Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu
                165                 170                 175

gtt tgt ttt aaa aaa cgt att gaa gct atc cca caa att gat aag tac      576
Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr
            180                 185                 190

ttg aaa tcc agc aag tat ata gca tgg cct ttg cag ggc tgg caa gcc      624
Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala
        195                 200                 205
```

```
acg ttt ggt ggt ggc gac cat cct cca aaa tcg gat ctg gaa gtt ctg      672
Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Glu Val Leu
    210                 215                 220

ttc cag ggg ccc ctg                                              687
Phe Gln Gly Pro Leu
225
```

<210> 30
<211> 229
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
GST coding sequence derived from pGEX6p-1

<400> 30

```
Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro
1               5                  10                  15

Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu
            20                  25                  30

Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu
        35                  40                  45

Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys
    50                  55                  60

Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn
65                  70                  75                  80

Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu
                85                  90                  95

Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser
            100                 105                 110

Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu
        115                 120                 125

Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn
    130                 135                 140

Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp
145                 150                 155                 160
```

Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu
165 170 175

Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr
180 185 190

Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala
195 200 205

Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Glu Val Leu
210 215 220

Phe Gln Gly Pro Leu
225

<210> 31
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic protease cleavage site

<220>
<221> CDS
<222> (1)..(24)

<400> 31
ctg gaa gtt ctg ttc cag ggg ccc 24
Leu Glu Val Leu Phe Gln Gly Pro
1 5

<210> 32
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic protease cleavage site

<400> 32
Leu Glu Val Leu Phe Gln Gly Pro
1 5

<210> 33
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
nucleotide sequence that contains a Kozak signal,
start codon and NcoI-KpnI-XbaI-PstI linker

<400> 33
gatgcggtac caccatggtg tctagactgc ag 32

<210> 34
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
nucleotide sequence that contains an
AvrII-ApaI-SbfI linker

<400> 34
tgcctagggc cctgcagggt a 21

<210> 35
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
nucleotide sequence that contains an
SpeI-EcoRI-NsiI-NheI linker and stop codon

<400> 35
actagtgaat tcatgcattg agctagccat c 31

<210> 36
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
double stranded oligonucleotide linker

<220>
<221> modified_base
<222> (5)
<223> a, g, c, t, unknown or other

<220>
<221> modified_base
<222> (6)..(8)
<223> a, g, c, t, unknown or other and see
specification for further embodiments

<220>
<221> modified_base
<222> (9)
<223> a, g, c, t, unknown or other

<400> 36
ctagnnnnnt gca 13

<210> 37
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic nucleotide sequence that contains HindIII-BamHI sites

<400> 37
aagcttggaa ccggatcc 18

<210> 38
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic nucleotide sequence that contains BamHI and EcoRI restriction sites

<400> 38
ggatcctctt cagaattc 18

<210> 39
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic nucleotide sequence that contains c-myc epitope tag, stop codon and NheI restriction site

<400> 39
gagcagaaac tcatctctga agaggatctg tgagctagc 39

<210> 40
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
double stranded oligonucleotide linker

<220>
<221> modified_base
<222> (6)..(8)
<223> a, c, g, t, unknown or other and see
specification for further embodiments

<400> 40
agcttnnng 9

<210> 41
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
double stranded oligonucleotide linker

<220>
<221> modified_base
<222> (6)..(8)
<223> a, c, g, t, unknown or other and see
specification for further embodiments

<400> 41
aattcnnna 9